Computing Nonnegative Rank Factorizations

Stephen L. Campbell
Department of Mathematics
North Carolina State University
Raleigh, North Carolina 27650

and

George D. Poole*
Department of Mathematics
Emporia State University
Emporia, Kansas 66801

Submitted by Robert J. Plemmons

ABSTRACT

The existence of nonnegative generalized inverses in terms of nonnegative rank factorizations is considered. An algorithm is presented which computes a nonnegative rank factorization of a nonnegative matrix when a nonnegative 1-inverse exists.

I. INTRODUCTION

Suppose A is an $m\times n$ matrix with rank $r(A)=t$. The product FG is called a *rank factorization* of A if F is an $m\times t$ matrix of full rank, G is a $t\times n$ matrix of full rank, and $FG=A$. If both F and G are nonnegative, FG is called a *nonnegative rank factorization*.

Whenever they exist, such factorizations yield constructive procedures for finding various kinds of generalized inverses of A. For example, in [3] Theorem 1.3.2 and Theorem 7.8.2 illustrate how rank factorizations can be used effectively to obtain the Moore-Penrose inverse and Drazin inverse of A. Berman and Plemmons [1] have shown the usefulness of the rank factorization of a matrix in constructing generalized inverses and nonnegative generalized inverses of various kinds.

*This author's work was done while on sabbatical at North Carolina State University, 1979–1980.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/81/010175+8$02.50

X is called a *1-inverse* of A if $AXA = A$, and a *2-inverse* if $XAX = X$. If X is both a 1-inverse and 2-inverse of A, X is called a *reflexive inverse* of A. Berman and Plemmons [1] have shown that among nonnegative matrices, A has a nonnegative 1-inverse (or nonnegative reflexive inverse) if and only if A has a very special type of nonnegative rank factorization.

In this paper we shall give necessary and sufficient conditions for a nonnegative matrix to have a nonnegative rank factorization. Furthermore, we shall give an effective algorithm for determining whether or not A has a nonnegative reflexive or 1-inverse. When either inverse exists, the algorithm will determine the cone dimension of the rows of A and one nonnegative rank factorization of A.

X is called the *Drazin inverse* of A if for some k, $XAX = X$, $XA = AX$, and $XA^{k+1} = A^k$.

In characterizing matrices which have a nonnegative Drazin inverse A^D, Pye [5] has recently conjectured that if A and its Drazin inverse A^D are both nonnegative, then A necessarily has a nonnegative rank factorization. He observed that if A was nonnegative and A had a nonnegative *Cline factorization* [4] for which the reverse product of the last factorization is either zero or is a *monomial* (that is, the product of a positive diagonal matrix and a permutation matrix), then A^D was necessarily nonnegative. Of course, A^D could still be nonnegative without A having such a Cline factorization.

Suppose A is square and nonnegative. The existence of a nonnegative 1-inverse or nonnegative reflexive inverse for A guarantees the existence of a nonnegative rank factorization for A. Since $X = 0$ is a nonnegative 2-inverse for any A, the existence of a nonnegative 2-inverse does not.

A natural question, then, is what types of nonnegative 2-inverses, other than reflexive ones, guarantee the existence of nonnegative rank factorizations. One of the few naturally occurring 2-inverses that is not necessarily a reflexive inverse is the Drazin inverse.

We shall show that, in general, if A is nonnegative and the index of A exceeds one, then the nonnegative rank factorization question of A is independent of A^D being nonnegative. In particular, we shall exhibit a nonnegative matrix A for which A^D is nonnegative and for which no nonnegative rank factorization exists, thus showing that Pye's conjecture is false.

II. NILPOTENTS AND NONNEGATIVE RANK FACTORIZATIONS

It is known [7] that a nonnegative matrix B need not have a nonnegative rank factorization.

Let $N = \{v_1, v_2, \ldots, v_n\}$ denote a set of nonnegative vectors. $C(N)$ will denote the polyhedral cone generated by N under nonnegative linear combinations. $\dim C(N)$ will denote the *cone dimension* of $C(N)$, namely the number of edges of $C(N)$.

We shall now consider a sequence of results which will lead to the example promised above. Their proofs are either omitted or sketched.

PROPOSITION 2.1. *Suppose A is nonnegative, $r(A)=t$, and N denotes the set of rows of A. A has a nonnegative rank factorization if and only if there exists a set M of t nonnegative vectors such that $C(N)\subseteq C(M)$ and* $\dim C(M)=t$.

Proof. The proof is based on the fact that if FG is a nonnegative rank factorization of A, the rows of A must be nonnegative linear combinations of the rows of G. ∎

Since the cone dimension of A is greater than or equal to its rank, one may observe from Proposition 2.1 that any nonnegative matrix with cone dimension less than four has a nonnegative rank factorization. Furthermore, (as observed in [7]) any nonnegative matrix with rank less than three has a nonnegative rank factorization.

A geometric characterization similar to Proposition 2.1 was obtained by Thomas [7], and was recently generalized by Wall [8] to nonnegativity and nonnegative rank factorizations with respect to a cone. The following result also appears in [9].

PROPOSITION 2.2. *If A is nonnegative and nilpotent, then there exists a permutation matrix P such that PAP^T is strictly upper triangular.*

Proof. Suppose $A^k=0$. Let j be a vector of ones. Then $A^kj=0$ implies that A must have a row of zeros, which, by a permutation similarity, we may assume is the last row. But each principle submatrix of A must also be nilpotent and hence must also contain a zero row, which, as before, we can assume to be the last one. ∎

PROPOSITION 2.3. *Suppose B is a nonnegative matrix of order n and $r(B)=t$. Suppose A is the square matrix of order $m=2n$ where*

$$A=\begin{bmatrix} 0 & B \\ 0 & 0 \end{bmatrix}.$$

Then A hs a nonnegative rank factorization if and only if B has a nonnegative rank factorization.

EXAMPLE 2.1. Consider the matrix B of order four given below:

$$B=\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}. \tag{2.1}$$

The rank of B is three. If N denotes the rows of B, $\dim C(N)=4$. Since the edges of the cone $C(N)$ lie on four distinct faces of the nonnegative orthant cone C^+, no triangular cone (one with three edges) in C^+ can contain $C(N)$. Therefore, by Proposition 2.1 B has no nonnegative rank factorization. This is not the first example of such a matrix, but it does illustrate the use of Proposition 2.1. Furthermore, Proposition 2.1 shows why four is the order of the smallest such example (see [7]).

Now define A to be the matrix in Proposition 2.3 where $n=4$, $m=8$, and B is the matrix defined in (1) above. Then A is nonnegative and nilpotent. Hence, $A^D=0$ is nonnegative. According to Proposition 2.3, since B has no nonnegative rank factorization, neither does A.

III. ALGORITHM FOR NONNEGATIVE RANK FACTORIZATIONS

An important observation from Proposition 2.3 is that the characterization of nonnegative matrices having nonnegative rank factorizations is equivalent to the characterization of nonnegative nilpotent matrices having nonnegative rank factorizations. This observation together with Proposition 2.1 means that very little ground is gained in determining when a nonnegative matrix A has a nonnegative rank factorization by considering the mixed case apart from the nilpotent case (A is called *mixed* if A is neither nilpotent nor of index one).

Proposition 2.1 is very geometrical in nature, and it is not always easy to determine algebraically when one cone is contained in another cone of smaller dimension. Therefore, any improvement of Proposition 2.1 requires one to generate algebraic and combinatorial formulae which probably apply only to rather special cases. Therefore, it would be extremely useful to obtain an algorithm to determine when nonnegative rank factorizations exist. No such algorithm is contained in [7] or [8].

Berman and Plemmons [1] observed the important fact that if there is one nonnegative rank factorization $A=FG$ such that F,G contain monomials of rank t where $\operatorname{rank}(A)=t$, then all nonnegative rank factorizations have this property. Therefore, in determining whether A has a nonnegative 1-inverse or nonnegative reflexive inverse, it is sufficient to determine that A

has no nonnegative rank factorization, or for one such nonnegative rank factorization to determine whether the factors contain the appropriate monomial submatrices.

We shall now introduce an algorithm which determines a nonnegative rank factorization whenever a nonnegative 1-inverse exists. It is a modification of the method one would use to compute ordinary rank factorizations of A.

A *nonnegative elementary row operation* is one which either multiplies row i by a positive scalar α, interchanges row i and row j, or multiplies row i by a negative scalar $-\alpha$ and adds to row j.

The corresponding *reverse elementary column operations* are the multiplication of column i by $1/\alpha$, the interchanging of columns i and j, or the multiplication of column j by α and subsequent addition to column i.

Suppose A is a nonnegative m by n matrix. Let $\{r_1, r_2, \dots, r_m\}$ denote the set of rows of A. Let $z(r_i)$ denote the *zero set* of r_i, that is, the set of integers k for which the kth component of r_i is zero. Now row reduce A as follows.

If $z(r_i) \subseteq z(r_j)$, then by subtracting a nonnegative multiple of r_j from r_i we obtain a new nonnegative ith row $\hat{r}_i$ such that $z(r_i) \not\subseteq z(\hat{r}_i)$. If at any time a row is identically zero, move it to the bottom of A by interchanging rows. After a finite number of nonnegative elementary row operations have been performed, the matrix A has been row reduced as far as possible. That is, there is a sequence of t nonnegative elementary row operations and an $s \times n$ matrix R containing no zero rows such that

$$P_t P_{t-1} \cdots P_1 A = \begin{bmatrix} R \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{r}_1 \\ \vdots \\ \tilde{r}_s \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{3.1}$$

where $z(\tilde{r}_i) < n$ for all $i = 1, 2, \dots, s$, and for any $1 \leqslant i, j \leqslant s$, $z(\tilde{r}_i) \subseteq z(\tilde{r}_j)$ implies that $i = j$. From these observations we obtain

PROPOSITION 3.1. *If the nonnegative $m \times n$ matrix A has a nonnegative 1-inverse, the integer s in* (3.1) *is not only the dimension of the smallest polyhedral cone in C^+ that contains the polyhedral cone generated by the rows of A but also $r(A)$.*

Proposition 3.1 together with Proposition 2.1 show that if $s > r(A)$, no nonnegative rank factorization exists for A.

Suppose now that A has been row reduced to (3.1). Then

$$A=P_1^{-1}P_2^{-1}\cdots P_t^{-1}\begin{bmatrix} R_{s\times n} \\ 0 \end{bmatrix}=\begin{bmatrix} Q & T \end{bmatrix}\begin{bmatrix} R \\ 0 \end{bmatrix}, \tag{3.2}$$

where Q is $m\times s$. Since $A=QR$ and $r(R)=r(A)$, QR is a nonnegative rank factorization of A provided $s=r(R)$.

These observations can be organized into the following (computer adaptable) algorithm.

ALGORITHM. Suppose A is a nonnegative matrix of order $m\times n$. Construct the augmented matrix $B=[A \mid I]$ where I is $m\times m$. Perform a sequence of nonnegative elementary row operations P_i which row reduces A to the form given in (3.1). For each P_i, perform the corresponding reverse elementary column operation S_i to I in the same order. Thus B is transformed to

$$\tilde{B}=\left[\begin{array}{c|cc} R_{s\times n} & Q_{m\times s} & T \\ 0 & & \end{array}\right] \tag{3.3}$$

If $s>r(A)$, there is no nonnegative rank factorization. If $s=r(A)$, then QR is the desired nonnegative rank factorization of A.

Furthermore, it is a routine matter to run a row check on Q and a column check on R to determine if they contain monomial submatrices of rank s. If both R and Q contain monomial submatrices of rank $s=r(A)$, Theorem 4 of [1] applies.

To illustrate the algorithm consider the matrix

$$A=\begin{bmatrix} 4 & 0 & 4 & 11 \\ 0 & 2 & 2 & 0 \\ 4 & 0 & 4 & 8 \\ 5 & 0 & 5 & 15 \\ 1 & 1 & 2 & 2 \end{bmatrix} \tag{3.4}$$

We shall use αr_i, $r_i\leftrightarrow r_j$, $r_j-\alpha r_i$ to denote the nonnegative elementary row operations discussed above, and $(1/\alpha)c_i$, $c_i\leftrightarrow c_j$, $c_i+\alpha c_j$ to denote the corresponding reverse elementary column operations.

The reader may decide that there is a preferred "pivot row" (such as one with the most zeros) which should be used to begin the "sweepout" process. However, we shall start with the third row, by interchanging it with the first row, and compare its zero set with the zero sets of the other rows, performing row operations where allowed.

Our first step is to augment A with an identity matrix of order 5:

$$\left[\begin{array}{cccc|ccccc} 4 & 0 & 4 & 11 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 & 1 & 0 & 0 & 0 \\ 4 & 0 & 4 & 8 & 0 & 0 & 1 & 0 & 0 \\ 5 & 0 & 5 & 15 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 2 & 2 & 0 & 0 & 0 & 0 & 1 \end{array}\right]. \tag{3.5}$$

Perform $r_1 \leftrightarrow r_3$ and $c_1 \leftrightarrow c_3$ followed by $\frac{1}{4}r_1$ and $4c_1$ to get

$$\left[\begin{array}{cccc|ccccc} 1 & 0 & 1 & 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 & 1 & 0 & 0 & 0 \\ 4 & 0 & 4 & 11 & 4 & 0 & 0 & 0 & 0 \\ 5 & 0 & 5 & 15 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 2 & 2 & 0 & 0 & 0 & 0 & 1 \end{array}\right]. \tag{3.6}$$

Now $z(r_2) \not\subseteq z(r_1)$, while $z(r_3) \subseteq z(r_1)$, $z(r_4) \subseteq z(r_1)$, and $z(r_5) \subseteq z(r_1)$. Perform first the operations $r_3 - 4r_1$, $c_1 + 4c_3$, then $r_4 - 5r_1$, $c_1 + 5c_4$, and finally $r_5 - r_1$, $c_1 + c_5$. The matrix (3.6) is now

$$\left[\begin{array}{cccc|ccccc} 1 & 0 & 1 & 2 & 4 & 0 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 5 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]. \tag{3.7}$$

Next consider r_2. $z(r_5) \subseteq z(r_2)$, while $z(r_i) \not\subseteq z(r_2)$ if i is neither 2 nor 5. Perform $\frac{1}{2}r_2$, $2c_2$ and then $r_5 - r_2$, $c_2 + c_5$. The matrix (3.7) is now

$$\left[\begin{array}{cccc|ccccc} 1 & 0 & 1 & 2 & 4 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 5 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right]. \tag{3.8}$$

Finally consider r_3. Note that $z(r_1) \subseteq z(r_3)$, $z(r_4) \subseteq z(r_3)$. Rescale r_3 by the operations $\frac{1}{3}r_3$, $3c_3$. Then perform $r_1 - 2r_3$, $c_3 + 2c_1$, followed by $r_4 - 5r_3$,

c_3+5c_4, so that (3.8) is now

$$\left[\begin{array}{cccc|ccccc} 1 & 0 & 1 & 0 & 4 & 0 & 11 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 & 15 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 2 & 0 & 1 \end{array}\right]. \tag{3.9}$$

Thus, (3.5) has been transformed to (3.9), which is the form required in (3.3). Since the cone dimension and rank of A agree,

$$Q=\begin{bmatrix} 4 & 0 & 11 \\ 0 & 2 & 0 \\ 4 & 0 & 8 \\ 5 & 0 & 15 \\ 1 & 1 & 2 \end{bmatrix}, \qquad R=\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.10}$$

are two factors of a nonnegative rank factorization for A.

Observe that R contains a monomial of order three while Q does not. Consequently, A does not have a nonnegative 1-inverse. In determining whether A has a nonnegative 1-inverse, it is easier to check for monomial submatrices in the two factors Q, R rather than check for a monomial submatrix in A.

REFERENCES

1 Abraham Berman and Robert J. Plemmons, Inverses of nonnegative matrices, *Linear and Multilinear Algebra* 2:161–172 (1974).

2 Abraham Berman and Robert J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic, 1979.

3 S. L. Campbell and C. D. Meyer, *Generalized Inverses of Linear Transformations*, Pitman, 1979.

4 R. E. Cline, Inverses of rank invariant powers of a matrix, *SIAM J. Numer. Anal.* 5:182–197 (1968).

5 Wallace C. Pye, Nonnegative Drazin inverses, *Linear Algebra and Appl.* 30:149–153 (1980).

6 Wallace Pye and Melvin Jeter, Nonnegative rank factorizations, unpublished communication.

7 L. B. Thomas, Rank factorizations of nonnegative matrices, *SIAM Rev.* 16:393–394 (1974).

8 J. R. Wall, Rank factorizations of positive operators, *Linear and Multilinear Algebra* 8:137–144 (1979).

9 S. K. Jain, V. K. Goel and E. K. Kwak, Nonnegative mth roots of nonnegative 0-symmetric idempotent matrices, *Linear Algebra and Appl.* 23:37–52 (1979).

Received October 1979; revised 15 February 1980